

== GEOGRAPHIC INFORMATION SYSTEMS ISSUES ==

PAST, PRESENT AND FUTURE OF GEOGRAPHIC INFORMATION SYSTEMS IN WATER RESOURCES

Richard M. Males, President

RMM Technical Services, Inc., Cincinnati, Ohio

Walter M. Grayman, Consulting Engineer

Cincinnati, Ohio

Introduction

Geographic Information System (GIS) technology is an increasingly popular methodological approach to working on a variety of problems of interest in the field of water resources. Many of the users of GIS view this technology as a relatively new tool with little history or applications experience in our field. In fact, GIS technology has been used extensively in a variety of water resource-related studies for the past two decades. The character of much recent work is simply 're-invention of the wheel', as researchers and workers proceed without examining the literature and learning from prior experience. The purpose of this paper is to provide some idea of the threads that have led to the current GIS technology and to propose some of the directions for improvement of GIS technology so that it is a better tool for water resources applications. The authors write from the perspective of having applied GIS and spatial data technologies to a variety of water resource projects and problems over the last twenty years.

Background of GIS Technology

GIS technology includes, at a minimum, the ability to enter, display, edit, and manipulate computer-stored information that has locational attributes. It differs from computer-aided mapping or drafting technologies in that, in addition, GIS technology provides the ability to perform spatial analysis on this data, i.e. to examine the data using concepts that are unique to locational data, such as nearness, sequential location on a network, etc.

Frequently, GIS technology stores topological relationships between spatial data elements to accomplish this end, where mapping or drafting technologies may be restricted to storage of geometry only. Finally, GIS technology may include varying degrees of modeling capability, that is, the ability to use the stored spatial data as the basis for analytical or simulation modeling of phenomena.

There are a variety of technical problems involved in storing and manipulating spatial data in a computer. Most current-day computers are, by nature, sequentially operating machines. Tasks that are very simple for a human being, such as looking at a map and describing what cities are near which rivers, is significantly more difficult for a computer, requiring, at minimum: a specific algorithmic definition of 'nearness'; methods for storing the boundaries and traces of cities and rivers, and distinguishing between the two; and enumeration methods to insure that the nearness algorithm is applied to all necessary cases. Real world spatial data is, further, inherently complex. Simplifications, such as assuming a one-dimensional, tree-structured river, may be adequate for some modeling applications, but fail to capture the reality of complex, braided, deltaic patterns. Data structures used to store spatial data represent compromises between simplicity, data storage density, and representation of 'reality'. To this date, no really good method of computer representation of sub-surface geologic structures has been developed.

GIS technology has evolved from a variety of sources, including computer mapping, survey-

ing and photogrammetry, and planning (in particular certain schools of landscape architecture). Computerized map-making has been of interest to cartographers since the '60s. The goal was simply to make the production and modification of maps simpler. The growth of such technology has paralleled the rise of computer-aided drafting techniques in general. Both computer-aided drafting and mapping initially had as their goal the generation of paper products, i.e. maps or drafting sheets. Eventually, it was recognized that the data base that is required to generate the paper products can be used for other analyses, and a more 'model-oriented' approach was born, in which the data stored was seen as a model of the world, rather than as simply the set of coordinates necessary to draw the desired map. Distinctions were made between cartographic applications, in which the maps needed to be of publication quality, and planning and analysis applications, for which computer generated maps could make use of the only commonly available output device, the line printer. The SYMAP program (Robertson, 1967), developed at Harvard University's Laboratory for Computer Graphics and Spatial Analysis in 1963, provided a mechanism for generating various forms of maps on a line printer, and was widely used.

Another source for GIS technology was computerized photogrammetry and surveying. Again in the early '60s, the desire to manipulate spatial data in a computer led to such well-known civil engineering programs as COGO, developed at MIT for calculating surveying analyses on coordinate data. At the same time, the concept of the Digital Terrain Model (DTM) was developed, in which the computer would be used to store a digital data base representing the earth's surface. Early attempts were made to use DTM's for analysis and modeling - one project at MIT in 1964 involved attempting to develop optimal highway routings based on modeling of costs of alternative routes using DTM's. (This project was ultimately unsuccessful due to constraints inherent in the grid-based technology that was used to store the digital terrain model - a sufficiently small-sized grid to adequately represent the terrain resulted in far too much data,

and a larger sized grid failed to provide the necessary terrain resolution).

Much of GIS modeling technology, as it is used today, is largely an outgrowth of planning approaches that are based on the work of Ian McHarg, as articulated in his book "Design with Nature" (McHarg, 1969). McHarg, a landscape architect (and strong critic of civil engineers), popularized, in the late '60s, a methodology of planning called 'suitability analysis'. His manual methods involved overlaying a grid on the area to be studied, and comparing and combining values for different types of attributes in a grid cell, to determine the 'suitability' of each grid cell for various uses. When computerized, this became the basis of the common 'overlay analysis' or 'weighting' of GIS technology, which served as the basic 'modeling' technology of GIS for many years.

Due to the two different GIS ancestors of cartography and planning, the technologies in GIS were divided into the 'polygon-based' approaches derived from cartography, in which detailed coordinates of the curving lines of maps are stored, and the 'grid-based' approach of the planners. The polygon approach provides good representation of mapped information, but proved difficult to use in analysis. The 'grid-based' approaches derived from planning, and from mapping programs based on the line printer such as SYMAP (Robertson, 1967), were constrained to grid-cell oriented representation. The grid approach, based on a matrix and thus easily handled by FORTRAN programming, provided easy analysis capabilities of the overlay type, but did not provide very good representations. In the early and mid-70's, the arguments about the merits of the two approaches were continual. Remote sensing data was becoming available, in grid (raster) format. Eventually, the problem was 'resolved' by the use of polygon-to-grid technology, in which data is captured originally in polygon format, and converted to a grid-based format for analysis (Comprehensive Planning Organization, 1974). Later development of more sophisticated and efficient algorithms allowed for 'polygon overlay', in which two separate

representations of an area, mapped as polygons (e.g. soil type and slope), could be combined to generate a third map showing the intersection of these two sets of polygons (in this case, used as a surrogate for septic tank suitability). An effective 'toolkit' for manipulating spatial data requires capabilities for handling both polygon and grid-based data, and transforming representations between the two types. (It should be noted that, in these transformations, data is either lost or artificially generated).

In general, terrain and network data, both of prime significance to water resources applications, were ignored. Terrain data presented a significant problem for incorporation into either the polygon or grid systems. The innate variation in terrain means that selection of grid size is critical, and it is essentially impossible to select an appropriate grid size for an area that captures fine-scale terrain variations in the steep areas without storing vast quantities of redundant data in the flatter areas. This became known as the 'variable grid resolution' problem, and a variety of methods were tried to overcome this, including the use of hierarchical grids (where sub-grids were used in detail areas) and triangulated surfaces (called TIN's, for Triangulated Irregular Networks) (Peucker, 1973; Grayman, 1975; Gold, 1976). The hierarchical grid method never proved workable in practice, but the TIN has developed into a fundamental method of storing terrain data in GIS, and has also become part of the 'toolkit'.

By and large, the issue of handling terrain data was considered to be a specialized field, not really in the mainstream of GIS concern. Accordingly, few systems made any attempt to adequately integrate terrain data with 'coverage' data such as soils, land use, etc. Network data was usually considered to be the province of transportation planning, and few applications were oriented towards stream, water, or sewer networks. Modeling within early GIS technology was, in general, restricted to weighting, distance measures, and overlay analysis within grid cells. There was little or no true simulation or design modeling, even though

such models were available outside the world of GIS. Only recently has the general GIS community given any attention to integrating GIS with complex modeling capability.

Water Resources Applications of GIS Technology

For applications in water resources, GIS technology in general must handle coverage data relating to land use, land cover, geology, and soils. Network-oriented data (streams, water distribution systems, sewer systems) and terrain information are very significant. Frequently, sub-surface information is important (e.g. for ground water modeling). In the field of water resources, the major spatial data base was EPA's STORET water quality data base (initially implemented by the US Public Health Service in 1964), storing stream quality information that could be referenced by location (primarily by river mile index or latitude-longitude). In the early days of STORET development, the system was crude and difficult to use, and spatial retrieval capability was minimal, but the system has since been expanded and made into a powerful tool with significant spatially-oriented data retrieval and display capabilities.

EPA also maintains the Reach File, a data system initiated in the 1970's, that is organized by hydrologic structure and has significant spatial data retrieval and analysis capabilities organized by hydrologic units (Horn and Grayman, 1992). Stream reaches within the system have been digitized, and a good deal of data (discharger location, water supply intake location, etc.) has been referenced to the reach. The system maintains reach connectivity, allowing for routing-type analyses, and planning-level modeling of flows, toxics discharges, etc. A major upgrade of the Reach File is currently underway which will result in the inclusion of all streams and waterbodies found on the 1:100000 scale USGS digital line graph (DLG) of the United States.

The US Army Corps of Engineers has had a long involvement with GIS technology, starting

with sponsorship of research on resource analysis methods (Harvard University Department of Landscape Architecture, 1969). The Honey Hill study using this methodology (1971) was followed by the Santa Ana River Basin Study (1975) using grid-based methods for river basin planning. Problems with the Santa Ana study led to designation of the Hydrologic Engineering Center (HEC) of the Corps as a center of expertise and development work for GIS work within the Corps. HEC worked with grid-based technology, and linked it to HEC models, in particular the HEC-1 model, developing the HEC-SAM (Spatial Analysis Methodology). HEC-SAM methods were tested on a Corps Flood Plain Information Study (Phase I Oconee Basin Pilot Study, 1975), and were later widely used in a variety of Corps studies. The study and approach are notable for the combination of GIS and hydrologic modeling technology, in particular because the modeling was not simplified or subordinated to the GIS aspects of the study - rather, the GIS served as a database to 'feed' the models used. HEC has continued to utilize spatial analysis technology (preferring the term spatial analysis to GIS) and has produced a variety of computerized tools that make use of spatial data bases. HEC is currently working to develop the next generation of these tools using GIS and database technology.

Studies of non-point source pollution, given great impetus by Public Law 92-500 in the early '70s, led to the development and use of a number of GIS systems. Most such systems used grid-based technologies and simple overlay analysis or routing as the modeling components (Beasley and Huggins, 1982). More sophisticated systems were developed, however, to provide enhanced modeling capability. In particular, the ADAPT (Areal Design and Planning Tool) system was a TIN-based GIS, developed in 1972, with the primary objective of serving as a GIS-based modeling system for engineering design of wastewater systems and hydrologic analyses. The ADAPT system (Grayman, 1975, Males, 1980) used the triangulated irregular network technique of storing terrain, but also used each triangular element as a grid cell, for storing soil, land use, and land cover

data. The edges of the triangles formed links of stream systems, or boundaries of polygons, providing an integrated network, cell, polygon, and digital terrain system. This system was adopted in the late 1970's by the Ohio Environmental Protection Agency as the basis of the statewide Planning and Engineering Management System of Ohio and is still in use by that agency. The ADAPT system was notable for having variable resolution cells (i.e. triangle size could vary, depending upon the level of detail that needed to be captured), for the integration of terrain, network, and coverage data, and for the manual nature of database creation - the triangular data structure was developed as a true surface model, and the triangular plates were hand-developed based on examination of topo maps. A range of hydrologic and non point source models were interfaced to the spatial data base including the MITCAT kinematic wave model with rainfall represented as a moving storm passing over the catchment. While eminently suitable for hydrologic modeling, the difficulties of manual data base development made this TIN/cell technology less appealing.

Present Status of GIS in Water Resources

The current popularity of GIS in the water resources field can be gaged by the large number of conference sessions (and full conferences) and technical papers related to GIS. For example, an issue of the ASCE Journal of Water Resources Planning and Management will be dedicated to the area of GIS in water resources later this year and a GIS oriented conference is planned by AWRA in 1993. At the recently held 'First International Conference/Workshop on Integrating Geographic Information Systems and Environmental Modeling' (Boulder, CO, September 1991), GIS applications in water resources were highlighted in many posters, demos and papers. In two papers presented at the conference, the present status of GIS and hydrologic modeling (Maidment, 1991) and GIS and land surface/subsurface modeling (Moore et al, 1991) were reviewed. Both papers emphasize that the current focus in the area of GIS and water resource modeling is in the interfacing of existing

models to commercially available GIS packages. Frequently this approach to 'shoehorning' sophisticated water resources models into existing systems (which in many cases are designed to accommodate transportation related applications) casts some doubt on the veracity of the resulting GIS-modeling system.

GIS in water resources is also being influenced by the related fields of computer aided drafting (CAD) and automated mapping-facilities management (AM/FM). All three fields share a common dependency on spatially arrayed data. Linkages between popular GIS and CAD are being built and packages that originally grew from one of the three areas are being expanded to include capabilities associated with the other areas.

Another area of development that incorporates the concepts of GIS is the field of spatial decision support systems (SDSS). SDSS combines GIS, data base management and modeling concepts generally on a workstation platform. Several examples of SDSS technology as applied to the water resources field were displayed at the conference in Boulder and an Engineering Foundation conference on SDSS in Water Resources is planned for later this year.

Most current applications of GIS technology in water resources start with a commercial GIS as a given, and then layer various tools/models on top of that. In this way, the limitations and capabilities of the selected GIS define the scope of the application. Given the desire for ease of use, and integration, this idea of a single GIS package is inherently appealing, but falls short in a number of ways. The user gains little knowledge of spatial manipulation techniques, and is at a loss to handle data that may not fit the capabilities of the GIS package. The internal manipulations (which may often result in significant smoothing, or artificial data creation, as in most contouring algorithms) are hidden from the user. Thus, given a set of points, which are then passed through a contouring algorithm, the user assumes that the contour map so generated is representative of the terrain, whereas

it may be far from adequate. Most GIS use is 'blind', and users are unaware of problems such as these. Further, given the wide variety of data sources, and the increasing availability of digital mapped data bases, some capability and understanding of how to access and transform spatial data in a variety of formats is important. Unfortunately, the academically and research-based groups that published algorithms and code as mechanisms for transferring GIS technology in the 1970's, such as the University of Michigan geography program exchange, the Harvard Laboratory for Computer Graphics, and the Oak Ridge National Laboratories, no longer fulfill this function there does not appear to be a significant impetus for the exchange of such basic GIS algorithms. At one time, there was available a large body of public domain (FORTRAN) code for a variety of spatial data manipulation functionalities - contouring, point-in-polygon, area of polygons, etc., etc. The availability of such an updated library today, preferably implemented in standard C so as to be functional across a variety of platforms, would be a great addition to the water resource researcher's toolkit, and to educators and students interested in exploring and extending spatial data manipulation technology.

Future Directions for GIS

The field of GIS can be compared to a speeding train. It is moving fast and definitely making progress. However, some pressure must be applied by the water resources field to assure that the train ends up where we want it to go. Some specific recommendations for influencing GIS development are presented below:

- Teaching of the fundamental technologies and theories of computerized spatial manipulation to water resources students and professionals in university courses
- Increased research on data structures and manipulation methods that are of import to water resources applications

- Development of a 'toolkit' approach, as opposed to a blackbox approach, to using spatial data in water resources application, such that a variety of tools are available, and may be used to accomplish the desired end
- Increased focus on the modeling efforts, and less on the 'gee, we are using a GIS' aspects, of studies
- Definition of the basic hydrologic functions that must be supported by spatial data manipulation technologies (e.g. overland flow representation, linkages between land and stream based processes, etc.), and conduct of research to develop and refine the needed methods and data structures;
- Re-focusing of hydrologically based models from lumped parameter models to distributed models, that reflect the greater availability of spatial data in GIS.
- Generating greater concern for issues of data quality and data representation in support of modeling.

Summary

Many of the elements of computerized spatial analysis have been used in water resources since the mid '60s, and sophisticated and complex systems incorporating GIS and modeling technology were developed, tested, and applied to a wide variety of projects starting in the early '70s. The findings of the early developers, that GIS technology must support complex engineering modeling, and that it must provide representations of terrain and network information, as well as 2-dimensional information (soils, land cover, etc.), for it to be of use in water resources applications, are still valid today, and need not be re-discovered by those just

entering the field. In the future, if the GIS field is to develop in a direction that best suits the needs of the water resources field, water resource professionals must take an active role to define the needs of the field and then to pressure developers to support these needs. Furthermore, the water resources field must internally support research and development of models which can more effectively take advantage of the greatly increased spatial data that is available through GIS technology.

References

- Beasley, D.B. and L. F. Huggins. 1982. "ANSWERS: User's Manual." Chicago, IL: U.S. EPA.
- Comprehensive Planning Organization. 1974. PIOS II Polygon to Grid Conversion, San Diego.
- Gold, C.M. 1976. "Triangular Element Data Structures." University of Alberta Computing Services, Users Applications Symposium Proceedings, Edmonton.
- Grayman, W. and R. M. Males, et al. 1975. "Land-Based Modeling System for Water Quality Management Studies." *Journal of the Hydraulics Division*, ASCE, Vol 101, No. HY5, May.
- Horn, C.R., and W. M. Grayman. 1992. "Water Quality Modeling with the EPA Reach File System." Submitted to the ASCE J.WRPMD.
- The Hydrologic Engineering Center. 1975. "Phase I Oconee Basin Pilot Study, Trail Creek Test." Corps of Engineers, September.
- IWR Report 71-9. 1971. "Honey Hill: A Systems Analysis for Planning the Multiple Use of Controlled Water Areas." Corps of Engineers, October.
- IWR Contract Report 75-3. 1975. "The Santa Ana Basin: An Example of the Use of Computer Graphics in Regional Plan Evaluation." Corps of Engineers, June.
- Maidment, D.R. 1991. "GIS and Hydrologic Modeling." First International Conference/Workshop on Integrating Geographic Information Systems and Environmental Modeling. Boulder, CO, September.
- Males, R.M. 1980. "ADAPT: The Use of Complex Models with a Geographic Data Base." Proceedings, Specialty Conference, The Planning and Engineering

Interface with a Modernized Land Data System,
ASCE.

McHarg, Ian. 1969. *Design With Nature*. Garden City, NY:
Natural History Press.

Moore, I.D., A.K. Turner, J. P. Wilson, S. K. Jenson, and
L. E. Band. 1991. "GIS and Land Surface-Subsur-
face Process Modelling." First International Con-
ference/Workshop on Integrating Geographic Infor-
mation Systems and Environmental Modeling,
Boulder, CO, September.

Peucker, T.K. 1973. "Geographic Data Structures Report
After Year One." Office of Naval Research Con-
tract N00014-73-C-0109, Department of Geogra-
phy, Simon Fraser University, Burnaby, B.C.

Robertson, J.C. 1967. "The SYMAP program for Computer
Mapping." *The Cartographic Journal*, Vol. 4, No.
2, pp. 108-113.

Steinitz, et al. 1969. "A Comparative Study of Resource
Analysis Methods." Department of Landscape Ar-
chitecture Research Office, Graduate School of
Design, Harvard University, July.